REPORT DOCUMENTATION

AD-A253 235

OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		FINAL 01 Nov 88 TO 30 Apr 92

4. TITLE AND SUBTITLE
VISUAL PERCEPTION OF STRUCTURE FROM MOTION

5. FUNDING NUMBERS
G AFOSR-89-0016
PE 61102F
PR 2313
TA A5

6. AUTHOR(S)
Dr James T. Todd

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Boston University
Center for Adaptive Systems
111 Cummington Street, Second Floor
Boston, MA   02215

AFOSR·T

8. PERFORMING ORGANIZATION REPORT NUMBER

3

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dr John F. Tangney
AFOSR/NL
Building 410
Bolling AFB DC   20332-6448

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

DTIC
ELECTE
JUL 24 1992
S
B
D

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release;
distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The research performed in this project has examined the abilities of human observers
to perceive 3D form from different types of optical structure within moving or
stationary visual images. The research has been organized into four general
problem areas, including the low level detection of coherent motion, the analysis
of 3D form from motion, the analysis of image shading and texture, and the
identification of image contours. Our basic strategy in all of these areas has
been to identify the key assumptions of current computational models; to test the
psychological validity of those assumptions using appropriate psychophysical
procedures; and, based on the results of those experiments, to develop alternative
models that more closely match the perceptual capabilities of actual human observers.
In contrast to most common amethods of 3D image analysis, which are designed to
compute precise metrical descriptions, our results have shown that human perception
is primarily concerned with more abstract aspects of object structure, such as
affine or ordinal properties, which are easier to computer and are more robust to
uncontrolled changes in viewing conditions.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
(U)	(U)	(U)	(U)

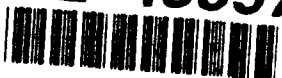Final Technical Report
Contract AFOSR-89-0016

The Visual Perception of Structure from Motion

Nov1, 1988 - April 30, 1992

Principal Investigator: James T. Todd

Boston University
Center for Adaptive Systems
111 Cummington St.
Boston, MA 02215

92 7 23 007

The research performed in my laboratory during the past three years under AFOSR support has been designed to investigate the fundamental mechanisms involved in the perceptual analysis of 3-dimensional form and motion. Much of this research has been specifically motivated from a computational perspective. Our basic research strategy has been to identify the key assumptions of various competing models; to empirically investigate the relative psychological validity of those assumptions using appropriate psychophysical procedures; and, when necessary, to develop alternative models based on revised assumptions that more closely match the perceptual capabilities of actual human observers.

There are several different properties of optical structure from which observers are able to obtain information about the structure of the environment. We have tried in our research to consider a broad spectrum of issues involving the analysis and integration of these different sources of information, and to identify any common mechanisms or strategies that may be involved in multiple domains of perceptual processing. A brief summary of these ongoing programs are described briefly below.

The Detection of Motion

One line of research performed in my laboratory has been designed to examine the competitive and coorperative interactions involved in the detection of coherent motion. For example, Mingolla, Todd & Norman (1992) have recently investigated how a field of moving contours each contained within its own small circular aperature can be spatially integrated to produce a perception of globally coherent motion. Previous theoretical analyses have shown that for translatory motion it is possible to determine the correct direction of translation using an intersection of constraints in velocity space. Our results suggest, however, that human observers analyze such patterns using a simple vector average of the component velocities perpendicular to each contour. We have also obtained evidence that the vector averages in different regions can be compared at a higher level of analysis to identify different categories of motion such as translation, rotation or divergence.

A related series of experiments by Kramer & Todd (1990) and Todd, Kramer & Norman (1992) has examined how processes of spatiotemporal integration can be used to filter spurious correlations (i.e., false targets) that can arise from the motions of densely textured patterns. In these studies we measured the maximum displacement thresholds for for moving patterns that varied in size, shape, duration and eccentricity, and contained variable amounts of uncorrelated noise. The results indicate that the human visual system may have multiple mechanisms of noise filtering to facilitate the detection of coherent motion, including averaging over space and correlations of features or motions over multiple time intervals.

The Analysis of Structure from Motion

A second line of research performed in my laboratory has been designed to investigate the visual perception of 3-dimensional structure from motion. During the past decade there have been numerous theoretical analyses proposed in the literature for computing an object's 3D form from a sequence of projected images. One of the primary results of these analyses is that a minimum of three distinct views is required in order to obtain an unambiguous

interpretation of an arbitrary configuration rotating in depth under orthographic projection. In an effort to test the psychological validity of this conclusion, we have performed numerous experiments using a wide variety of stimuli and response tasks (see Todd et al, 1988; Todd & Bressan, 1990; Todd & Norman, 1991; and Norman & Todd, 1992,1993). The results indicate that different aspects of 3D structure are judged with varying degrees of accuracy, but that none of these tasks are significantly influenced by increasing the number of distinct frames in an apparent motion sequence beyond two, if other factors are held constant.

Based on these findings, Todd & Bressan (1990) performed a mathematical analysis to determine what type of information is available from a 2-frame apparent motion sequence. This analysis showed that an object's structure can be specified from such displays up to an affine stretching transformation along the line of sight. That is to say, although there is not sufficient information in two views to specify a unique euclidean structure, the set of possible rigid interpretations is constrained to a one parameter family of structures that are related to one another by an affine stretching transformation. This analysis is also consistent with the psychophysical evidence from actual human observers.. Tasks that are theoretically possible based solely on an analysis of affine structure (e.g., object discriminations) invariably produce high levels of performance, whereas those that are theoretically impossible based solely on affine structure (e.g., 3D angle discriminations) invariably produce low levels of performance.

The Analysis of Shading and Texture

A third line of research performed in my laboratory has been designed to investigate the visual perception of 3D shape from shading and texture. The patterns of shading and texture in an image have several environmental determinants including the shape of the depicted object, its surface reflectance properties, and its pattern of illumination. Previous theoretical analyses have demonstrated that it is possible to compute the local orientation of any given surface region if provided with prior knowledge about the surface reflectance and illumination in that region. One of the goals of our research has been to psychophysically investigate whether similar prior knowledge is required for actual human perception. Our results show clearly that it is not. For example, Mingolla & Todd (1989) and Todd & Reichel (1989) demonstrated that observers' judgments of local orientation or ordinal depth for ellipsoid surfaces are largely unaffected by changes in surface reflectance or the direction of illumination as would be reasonable to expect based on current theory. Similarly, in the analysis of shape from texture or surface contours, Todd & Akerstrom (1987) and Todd & Reichel (1990) have shown that human perception does not require specific assumptions about the shapes of individual texture elements or how the contours align with the local surface curvature.

Much of our most recent research in this area has focussed primarily on the particular properties of surface structure that are perceptually specified by patterns shading and texture. This work has proceeded in an analogous fashion to our research on perceived structure from motion, in that we have designed a wide variety of response tasks to probe different aspects of local surface structure, including ordinal structure (Todd & Reichel, 1989; Reichel & Todd, 1990), affine structure (Reichel & Todd, 1991, 1992), and conformal structure (Mingolla & Todd, 1989). We have also performed numerous experiments to identify the processing mechanisms with which these different properties are perceptually analyzed. These include both masking and reaction time paradigms to asses the build-up of perceptual knowledge over time, and a delayed partial report paradigm to measure its decay in short term memory.

Contour Identification

Almost all computational analyses for determining 3D structure from visual images are designed to operate on abrupt changes in image structure over space called optical contours. A serious theoretical problem that is seldom mentioned in this context is that optical contours can arise due to a wide variety of physical phenomena (e.g., smooth occlusions, specular highlights, or abrupt changes in surface orientation, reflectance or illumination). It is important to keep in mind, however, that most computational analyses can only be used with a particular contour type. Thus, since contour identification seems to be a necessary prerequisite for the computational analysis of 3D form, Todd & Schnittman (1991) have recently performed a series of experiments to measure the minimal amount of contextual information needed to identify different types of contours in digitized images of natural objects. The results demonstrate the observers can reliably identify an optical contour in a few hundred milliseconds even when there is insufficient information in a scene to reliably identify any given object. We are currently attempting to analyze these reduced context images in an effort to discover which specific optical properties are used to discriminate between different types of contours.

Another important theoretical issue that needs to be considered in this context is that different types of contours produce different types of optical deformation when objects are viewed stereoscopically or in motion. Existing models for the computational analysis of structure from motion or stereopsis require that observed objects contain identifiable landmarks that can be tracked over time. There are many other types of commonly observed events, however, such as the deformations of shadows or smooth occlusion contours, for which these analyses are inadequate -- even as a local approximation. Human observers, in contrast, are able to interpret these events correctly. Todd, Norman & Fukuda (1992) have demonstrated that observers can reliably discriminate rigid from nonrigid motion when all that is visible is the optical deformation of its silhouette or its cast shadow on a background surface. A primary focus of our ongoing research is to discover how this is accomplished.

Publications

Todd, J.T., Akerstrom, R.A., Reichel, F.D., & Hayes, W. (1988). Apparent rotation in 3-dimensional space: Effects of temporal, spatial and structural factors. Perception & Psychophysics, 43, 179-188.

Akerstrom, R.A., & Todd, J.T. (1988). Perception of stereoscopic transparency. Perception & Psychophysics, 44, 421-432.

Todd, J.T. (1988). Perceived 3D structure from 2-frame apparent motion. Investigative Opthamology and Visual Science, 29, 265.

Todd, J.T. (1988). Minimal conditions for the perception of structure from motion. Bulletin of the Psychonomic Society, 26, 476-488.

Todd, J.T. (1989). Models of static form perception. In J.I. Elkind, S.K. Carol, J. Hochberg and B.M. Huey (Eds.), Human Performance Models for Computer-Aided Engineering (pp. 75-88). Washington, D.C.: National Academy Press.

Todd, J.T., & Reichel, F.D. (1989). Ordinal structure in the visual perception and cognition of smoothly curved surfaces. Psychological Review, 96, 643-657.

Mingolla, E., & Todd, J.T. (1989). Perception of solid shape from shading. In B.K.P. Horn & M.J. Brooks (Eds.), Shape from shading (pp. 402-442). Cambridge, MA: MIT Press.

Reichel, F.D., & Todd, J.T. (1989). Perceptual depth inversion of smoothly curved surfaces. Investigative Opthamology and Visual Science, 30, 252.

Todd, J.T. (1989). The short-range limit increases for multiple length apparent motion sequences. Bulletin of the Psychonomic Society, 27, 508.

Braunstein, M., & Todd, J.T. (1990). On the perceptual distinction between artifacts and information. Journal of Experimental Psychology: Human Perception and Performance, 16, 211-216.

Reichel, F.D., & Todd, J.T. (1990). Perceived depth inversion of smoothly curved surfaces due to image orientation. Journal of Experimental Psychology: Human Perception and Performance, 16, 653-664.

Todd, J.T., & Reichel, F.D. (1990). The visual perception of smoothly curved surfaces from double projected contour patterns. Journal of Experimental Psychology: Human Perception and Performance, 16, 665-674.

Todd, J.T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. Perception & Psychophysics, 48, 419-430.

Kramer, S.J., & Todd, J.T. (1990). Spatio-temporal integration in short-range apparent motion. Investigative Ophthalmology and Visual Science, 31, 519.

Reichel, F.D., & Todd, J.T. (1990). Visual knowledge of three dimensional surface structure. Investigative Ophthalmology and Visual Science, 31, 526.

Todd, J.T., & Mingolla, E. (1990). The visual perception of globally coherent motion. Bulletin of the Psychonomic Society, 28, 496.

Todd, J.T., & Norman, J.F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. Perception & Psychophysics, 50, 509-523.

Mingolla, E., Todd, J.T., & Norman, J.F. (1991). The perception of globally coherent motion. Investigative Ophthalmology & Visual Science, 32, 957.

Norman, J.F., & Todd, J.T. (1991). The perception of rigid motion from affine distortions in depth. Investigative Ophthalmology & Visual Science, 32, 958.

Reichel, F.D., & Todd, J.T. (1991). Metric knowledge of three dimensional surface structure. Investigative Ophthalmology & Visual Science, 32, 1179.

Todd, J.T., & Schnittman, P. (1991). The perceptual identification of image contours. Bulletin of the Psychonomic Society, 29, 529.

Norman, J.F., & Todd, J.T. (1992). The visual perception of 3-dimensional form. In G.A. Carpenter & S. Grossberg (Eds.), Neural networks for vision and image processing. Cambridge, MA: MIT press. pp. 93-110.

Mingolla, E., Norman, J.F., & Todd, J.T. (1992). Contextual information for the perception of surface gloss on curved surfaces. Investigative Ophthalmology & Visual Science, 33, 957.

Norman, J.F., Todd, J.T., & Fukuda, H. (1992). Perceived 3_d structure from boundary and shadow information. Investigative Ophthalmology & Visual Science, 33, 706.

Todd, J.T., Kramer, S.J., & Norman, J.F. (1992). Spatiotemporal integration in the detection of coherent motion. Investigative Ophthalmology & Visual Science, 33, 1050.

Mingolla, E., Todd, J.T., & Norman, J.F. (1992). The perception of globally coherent motion. Vision Research, in press.

Reichel, F. & Todd, J. T. (1992). Metric knowledge of three-dimensional surface structure. Society for Information Display Digest, in press.

Todd, J.T. (1993) On the optic sphere theory and the nature of visual information. In G. Jansson & S. S. Bergstrom (Eds.), Perceiving ecents and objects: A review of Gunnar Johansson's Research with commentaries. Hillsdale, NJ: Erlbaum, in press

Norman, J.F., & Todd, J.T. (1993). Visual perception of 3-dimensional affine stretching transformations. Perception & Psychophysics, submitted.